A METHODOLOGY FOR EVALUATING THE ACCURACY OF WAVE FIELD RENDERING TECHNIQUES

A.Canclini[§]*, P.Annibale*[†]*, F.Antonacci*[§]*, A.Sarti*[§]*, R.Rabenstein*[†]*, S.Tubaro*[§]

[§] Dipartimento di Elettronica ed Informazione
Politecnico di Milano, 20133 Milano, Italy
[†] University Erlangen-Nuremberg, Chair of Multimedia Communications and Signal Processing, LMS
Cauerstr. 7, 91058 Erlangen, Germany

ABSTRACT

In this paper we propose a methodology for assessing the accuracy of techniques of wave field rendering through loudspeaker arrays. In order to measure the rendered wave field we adopt a solution based on a circular harmonic analysis of the sound field captured by a virtual microphone array. As a result of this analysis stage, we are able to compare the target, the theoretical and the measured wave fields, which may differ due to the non-ideality in the loudspeaker array or in the environment that generates some spurious reverberations. Moreover, in order to quantify the error between target, theoretical and measured wave fields, we define some evaluation metrics, based on RMSE and modal analysis of the acquired wave fields. We show some experimental results on real data.

Index Terms— Loudspeaker array, virtual microphone array, wave field rendering evaluation

1. INTRODUCTION

Sound field reproduction methodologies, such as Wave Field Synthesis [1] and Higher-Order Ambisonics [2] are generally studied and developed under specific assumptions. The typical scenario consists of a distribution of ideal loudspeakers mounted in a completely anechoic acoustic environment. The performance of these systems is therefore predicted in a theoretical fashion by modeling the wave propagation inside the same ideal scenario. When the rendering methodologies are implemented on a real reproduction system, however, the non-ideality of both loudspeakers and environment may alter the quality of the rendered wave field. In order to evaluate this degradation, a methodology for accurately measuring the reproduced wave field inside the listening area is required.

This task is not trivial, because it implies a space-time sampling of the wave field within a wide region. If the sampling is performed with a single microphone moving inside the listening area, the number of acquisitions required to obtain a sufficiently dense reconstruction of the wave field makes this solution impracticable. On the other hand, if a regular grid composed of many microphones is used for sampling the wave field, their presence significantly alters the measured wave field. Furthermore, one has to be aware of the risk that the error introduced by the measurement methodology exceeds the degradation of the rendered wave field caused by the non-ideality of the reproduction system.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 226007. The work has been carried out during the visit of the first author at LMS.

In this paper we consider a well-established technique for the measurement of a two-dimensional wave field. As presented in [3] the wave field is sampled over a circle by means of a pair of rotating microphones. The wave field inside and outside the circle is extrapolated exploiting the Circular Harmonic Decomposition [4]. In order to compare target, theoretical and measured wave fields using synthetic parameters, we define some evaluation metrics. More specifically, we first introduce two RMSE metrics based on the difference between the theoretical and the measured wave fields. Although useful, these metrics are not able to distinguish at which propagation modes the error between the measured and theoretical wave field is introduced. Therefore, we define two other evaluation metrics that rely on a modal analysis based on the two-dimensional discrete Fourier transform. In this paper we are not aimed at a full validation of the measurement system; rather we show the feasibility of the methodology. More specifically, we compare different rendering techniques and show that the measurement methodology is able to capture the subtle differences between the measured and the theoretical wave fields that exist even in an ideal configuration (quasi-anechoic environment and high-quality loudspeakers).

The paper is organized as follows: Section 2 formulates the problem. Section 3 gives an overview of the technique for measuring a wave field presented in [3]. Section 4 describes the metrics used to evaluate the accuracy of the measured and theoretical wave fields. Section 5 shows the experimental setup while in Section 6 some results are presented. Finally, Section 7 draws some conclusions.

2. PROBLEM FORMULATION

A generic rendering system, shown in Fig. 1, is composed by an arbitrary distribution of M loudspeakers located at $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_M$. The

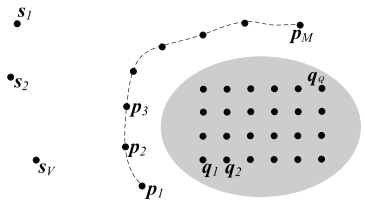


Fig. 1. A very general model of a rendering system.

goal of this system is to reproduce the wave field generated by a set of virtual point sources located at $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_V$ by means of the loudspeakers. The rendering system is designed for reproducing the target wave field inside the listening area, denoted by the grey-shaded

region shown in Fig. 1. We define an evaluation region composed of a set of Q points $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_Q$ regularly distributed inside the listening area. Let $S(\omega, \boldsymbol{q}_i)$ and $S^{\mathrm{T}}(\omega, \boldsymbol{q}_i)$ be the Fourier transforms of the target and theoretical wave fields at the point $\boldsymbol{q}_i$, respectively. Using the wavefield measuring technique described in Section 3 we acquire the measured wave field $S^{\mathrm{M}}(\omega, \boldsymbol{q}_i)$.

Our goal is to introduce some methodologies to evaluate the differences between $S(\omega, \boldsymbol{q}_i)$, $S^{\mathrm{T}}(\omega, \boldsymbol{q}_i)$ and $S^{\mathrm{M}}(\omega, \boldsymbol{q}_i)$.

3. MEASUREMENT OF THE WAVEFIELD

In this Section we summarize the methodology for the measurement of the wavefield presented in [3]. Adopting a polar coordinate system (radius ρ, angle ϕ), the time-domain Fourier transform of the sound pressure can be written as an angular Fourier series, whose coefficient can be further decomposed into circular harmonics

$$\begin{aligned} P(\omega, \rho, \phi) &= \sum_{\mu=-\infty}^{\infty} \mathring{P}_\mu(\omega, \rho) e^{j\mu\phi} \\ &= \sum_{\mu=-\infty}^{\infty} C_\mu(\omega) J_\mu(k\rho) e^{j\mu\phi} \ , \end{aligned} \tag{1}$$

where $C_\mu(\omega)$ represents the μth circular harmonic at frequency ω; $k = \omega/c$ is the wave number, c being the sound speed; J_μ is the Bessel function of the first kind and order μ. We notice that the circular harmonic coefficients depend only on the angular mode μ and the frequency ω. The knowledge of the sound pressure $P(\omega, \rho_0, \phi)$ on a circle with radius ρ_0 is therefore, in principle, sufficient to determine the wave field $P(\omega, \rho, \phi)$ inside and outside the circle, by adopting the following equalization:

$$C_\mu(\omega) = \frac{1}{J_\mu(k\rho_0)} \mathring{P}_\mu(\omega, \rho_0) \ . \tag{2}$$

However, the zeros of the Bessel functions at the denominator of eq.(2) make the direct implementation of this procedure impracticable. This issue is faced in [3] with the knowledge of the pressure and its gradient on the circle ρ_0. The combination of the signals acquired by the two sensors leads to a modified version of eq.(2), whose associated equalization function is free of zeros [3]. Moreover, the measurement of the wave field over the circle with radius ρ_0 in infinite positions is not feasible. A spatial sampling of $P(\omega, \phi)$ over the circle is therefore adopted in [3] through the rotating rig shown in Figure 2. The radius of the circle described by the arm is $\rho_0 = 0.74$ cm. The position of the arm is controlled by a stepper motor that, for every complete rotation, stops at N intermediate positions, under the assumption that the wave field is stationary during the rotation of the arm. This solution allows to sample the measurement circle with high angular resolution (up to 4000 positions can be used), keeping thus moderate the effect of modal aliasing introduced by the spatial sampling. Finally, we notice that the wave field $P(\omega, \rho_0, \phi)$ is measured on a circle with radius ρ_0, while the listening points $\boldsymbol{q}_i$ are located in arbitrary positions. Therefore an interpolation or extrapolation as described by eqs. (1) and (2) is required to map $P(\omega, \rho, \phi)$ on the listening points $\boldsymbol{q}_i$. The final result of the measurement procedure is the measured wave field $S^{\mathrm{M}}(\omega, \boldsymbol{q}_i)$.

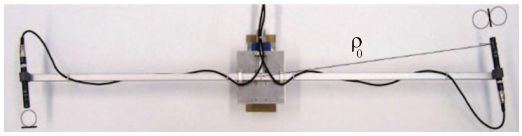


Fig. 2. The virtual array. An omnidirectional and a figure-of-eight microphones are used to emulate a cardioid microphone, as described in [3].

4. EVALUATION METHODOLOGY

In this Section we introduce the metrics used to evaluate the accuracy of the rendering technique. The evaluation methodology proposed here distinguishes between the target, theoretical and the measured wave field. The target wave field is the desired outcome of the reproduction (e.g. a plane wave). The theoretical wave field results from an approximation of the target wave field with a finite number of ideal loudspeakers and a given rendering technique. The measured wave field is the one resulting from acoustic measurements in a real-world rendering situation. First we introduce an evaluation metric that synthetically describes the accuracy of the theoretical and measured wave fields with respect to the target one. Later on we define a new evaluation metric that gives further insight into the evaluation of the accuracy, by performing a modal analysis on target, theoretical and measured wave fields.

4.1. Root mean square error metrics

As a first step we normalize the target wave field $S(\omega, \boldsymbol{q}_i)$ in order to make its energy unity. The resulting normalization coefficient $W(\omega)$ is used also to normalize the theoretical wave field. The resulting normalized wave fields are

$$\bar{S}(\omega, \boldsymbol{q}_i) = \frac{S(\omega, \boldsymbol{q}_i)}{W(\omega)} \ , \bar{S}^{\mathrm{T}}(\omega, \boldsymbol{q}_i) = \frac{S^{\mathrm{T}}(\omega, \boldsymbol{q}_i)}{W(\omega)} \ , \tag{3}$$

where $W(\omega) = \sqrt{\frac{1}{Q}\sum_{i=1}^{Q} |S(\omega, \boldsymbol{q}_i)|^2}$. This way, the target wave field is forced to have unitary energy in the evaluation region, independently on the set of virtual sources to be reproduced. As a consequence, the results related to different distributions of virtual sources can be directly compared.

Using the normalization in eq.(4.1), we introduce an evaluation metric based on the root mean square error between $\bar{S}(\omega, \boldsymbol{q}_i)$ and $\bar{S}^{\mathrm{T}}(\omega, \boldsymbol{q}_i)$.

$$e^{\mathrm{T}}_{\mathrm{mag}}(\omega) = \sqrt{\frac{1}{Q}\sum_{i=1}^{Q} |\bar{S}(\omega, \boldsymbol{q}_i) - \bar{S}^{\mathrm{T}}(\omega, \boldsymbol{q}_i)|^2} \ . \tag{4}$$

The metric in (4) synthetically quantifies the accuracy obtained by the theoretical wave field in reproducing the target one.

As far as the measured wave field is concerned, one could define an evaluation metric analogous to the one introduced for the theoretical evaluation. However, the measured wave field cannot be directly compared with the target wave field because its amplitude depends on the microphone gains and the loudspeaker volumes. Therefore, a normalization of the measured wave field is in order. More specifically, we propose to normalize it in such a way that the global energy within the evaluation region equals the energy of the theoretical wave field. In other words

$$\bar{S}^{\mathrm{M}}(\omega, \boldsymbol{q}_i) = \frac{S^{\mathrm{M}}(\omega, \boldsymbol{q}_i)\sqrt{\frac{1}{Q}\sum_{i=1}^{Q} |\bar{S}^{\mathrm{T}}(\omega, \boldsymbol{q}_i)|^2}}{\sqrt{\frac{1}{Q}\sum_{i=1}^{Q} |S(\omega, \boldsymbol{q}_i)|^2}} \ . \tag{5}$$

This way, the error between the theoretical and measured wave fields becomes

$$e^{\mathrm{M}}_{\mathrm{mag}}(\omega) = \sqrt{\frac{1}{Q}\sum_{i=1}^{Q} |\bar{S}(\omega, \boldsymbol{q}_i) - \bar{S}^{\mathrm{M}}(\omega, \boldsymbol{q}_i)|^2} \ . \tag{6}$$

4.2. Modal analysis

We notice that the evaluation metrics in eqs.(4) and (6) do not give insight on the distribution of the error for different wave fronts in the sound field. In order to address this issue, we perform a two-dimensional discrete Fourier transform on $\Re\{\bar{S}(\omega, \boldsymbol{q}_i)\}$, $\Re\{\bar{S}^{\mathrm{T}}(\omega, \boldsymbol{q}_i)\}$ and $\Re\{\bar{S}^{\mathrm{M}}(\omega, \boldsymbol{q}_i)\}$ to obtain $\hat{S}_{\mathrm{c}}(k_x, k_y)$, $\hat{S}_{\mathrm{c}}^{\mathrm{T}}(k_x, k_y)$ and $\hat{S}_{\mathrm{c}}^{\mathrm{M}}(k_x, k_y)$, respectively. The variables (k_x, k_y) are the spatial frequency bin coordinates. We omit the variable ω in the notation for the sake of compactness. A planar wavefront propagating in the environment is compactly described by the wavenumber k, inversely proportional to the wavelength, and by its direction of propagation φ. After some arithmetic we find that k and φ are related to k_x and k_y by

$$k = \tfrac{2\pi}{D}\sqrt{k_x^2 + k_y^2} \quad (7)$$

$$\varphi = \arctan k_y / k_x \ , \quad (8)$$

where D is the diameter of the area under analysis, in meters. For the sake of clarity, Figure 3 shows an example of target wave field on the left-hand side and the corresponding two-dimensional discrete Fourier transform on the right-hand side. We obtain therefore from

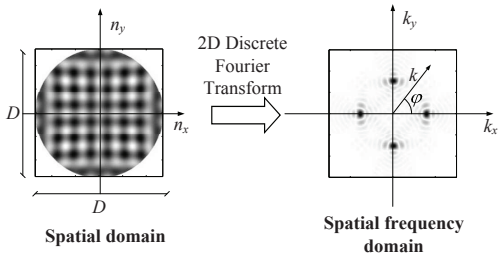


Fig. 3. An example of target wave field and the corresponding two-dimensional Fourier transform.

$\hat{S}_{\mathrm{c}}(k_x, k_y)$ the function $\hat{S}_{\mathrm{p}}(k, \varphi)$ using the change of variables in (7) and (8). We derive from $\hat{S}_{\mathrm{p}}(k, \varphi)$ two functions, $\hat{S}_1(k)$ and $\hat{S}_2(\varphi)$. The function $\hat{S}_1(k)$ aims at finding the distribution of the wavenumbers of the propagating wavefronts, regardless of their propagation direction and it is defined as

$$\hat{S}_1(k) = \sqrt{\sum_{\varphi} |\hat{S}(k, \varphi)|^2} \ . \quad (9)$$

In other words $\hat{S}_1(k)$ sums all the wavefronts in $\hat{S}_{\mathrm{p}}(k, \varphi)$, which are characterized by the same wavenumber. Under the hypothesis that the loudspeaker array is rendering a monochromatic signal, however, we obtain a wavenumber distribution characterized by a dominant wavenumber at which almost all the energy of the wavefront is contained. In order to distinguish wavefronts with the same wavenumber but with different directions of propagation, we define the function $\hat{S}_2(\varphi)$. The dominant wavenumber is found as

$$\hat{k} = \arg\max_k \hat{S}_1(k) \ . \quad (10)$$

The metric $\hat{S}_2(\varphi)$ assesses the distribution of the directions of propagation for the dominant wave number

$$\hat{S}_2(\varphi) = \hat{S}_{\mathrm{p}}(\hat{k}, \varphi) \ . \quad (11)$$

Adopting the same procedure described above, we derive $\hat{S}_1^{\mathrm{T}}(k)$ and $\hat{S}_2^{\mathrm{T}}(\varphi)$ from $\hat{S}^{\mathrm{T}}(k_x, k_y)$; and $\hat{S}_1^{\mathrm{M}}(k)$ and $\hat{S}_2^{\mathrm{M}}(\varphi)$ from $\hat{S}^{\mathrm{M}}(k_x, k_y)$. The comparison of the two functions allows us to analyze the differences for each wavenumber and, limited to the dominant wavenumber, for each direction of propagation between theoretical, measured and target wave fields.

5. EXPERIMENTAL SETUP

The loudspeaker array accommodates 48 high-quality emitters disposed on a circle with radius 1.5 m. The virtual microphone array has a radius of $\rho_0 = 0.74$ m and it samples the wave field on the circle at $N = 200$ intermediate positions. Signals are emitted and acquired at the frequency of $F_s = 44100$ Hz. Figure 4 shows the overall setup of the experiment. For each position, the signal is acquired for $T = 2$ s.

Fig. 4. The setup of the experiment.

In order to show the suitability of the measurement methodology and of the evaluation metrics described above for different rendering techniques, we perform the analysis for two rendering techniques, namely wave field synthesis (WFS, [1]) and geometric rendering (GR, [5]).

6. RESULTS

In this Section we present the comparisons between the target, the theoretical and the measured wave fields. In particular, we show two experiments.

The first experiment evaluates the accuracy of the measurement and of the reproduction of the wave field produced by the loudspeaker array when it simulates the presence of a monochromatic omnidirectional virtual source located at a distance d from the center of the array. We evaluated the accuracy of the two rendering systems for three distances $(3,\ 6,\ 10$ m) of the source. The source emits a monochromatic signal at 500 Hz. Table 1 and 2 show the theoretical and experimental evaluation results, respectively. In the experimental evaluation, we notice an improvement of the accuracy of the rendering systems as the distance of the source increases. In other words, the rendering of a source in the nearfield suffers from a distortion.

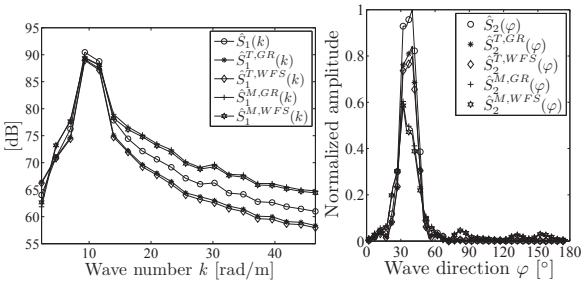
Table 1. Theoretical comparison between WFS and GR for a monochromatic wavefront.

d	$e_{\mathrm{mag}}^{\mathrm{T,WFS}}$	$e_{\mathrm{mag}}^{\mathrm{T,GR}}$
3 m	0.235	0.221
6 m	0.271	0.271
10 m	0.300	0.288

Table 2. Experimental comparison between WFS and GR.

d	$e_{\text{mag}}^{\text{M,WFS}}$	$e_{\text{mag}}^{\text{M,GR}}$
3 m	0.502	0.503
6 m	0.506	0.455
10 m	0.446	0.453

As far as the wave number and wave direction analyses are concerned, we show $\hat{S}_1(k), \hat{S}_2(\phi)$ for WFS and GR for the distance $d = 6$ m in Figure 5. We notice that the experimental wave number and wave direction distribution functions well fit the respective target distributions. This consideration applies for both WFS and GR.



(a) Wave-number distribution (b) Wave direction distribution

Fig. 5. Wave-number and wave direction distributions for the target, theoretical and measured wave fields of GR and WFS for the first experiment for $d = 6$ m.

For the second experiment we limit our theoretical and experimental evaluation to the GR technique. Here we aim at rendering the acoustics of a virtual environment. This is accomplished by spatially overlapping the acoustic beams generated by the image sources. The origin of the beams, their aperture and their orientation are computed by means of a fast beam tracing technique, presented in [6]. The environment we want to render is shown in Figure 6. The listening area is shown in black and it is encircled with red dots. The rendered sources are monochromatic at 500 Hz and at 1000 Hz. Table 3 shows the theoretical and experimental RMSE-based metrics. We notice that also in this case a degradation of the experimental results with respect to the theoretical ones is present, probably due to the spurious reverberations in the reproduction room.

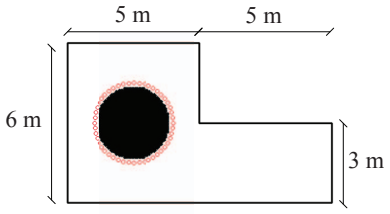
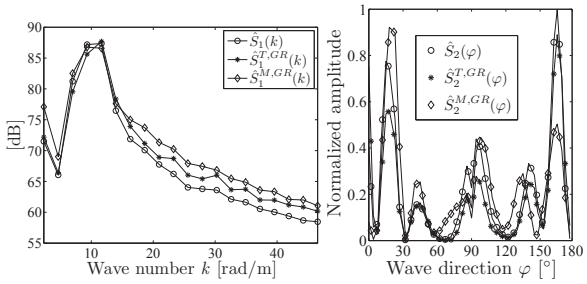


Fig. 6. Test environment used in the second experiment.

Table 3. Theoretical and experimental evaluation of GR for the simulation of the acoustics of the environment in Fig.6.

ω	$e_{\text{mag}}^{\text{T,GR}}$	$\hat{e}_{\text{mag}}^{\text{M,GR}}$
500 Hz	0.273	0.450
1000 Hz	0.282	0.540

Figure 7 shows the wave number and wave direction distributions for the second experiment for 500 Hz. We notice that the wave number distributions of theoretical and measured wave fields generally well fit the target one. The measurement methodology allows us to appreciate the presence of some spurious propagating wave fronts from in the range $(60^\circ \sim 90^\circ)$, which probably correspond to reverberations generated by the reproduction room.



(a) Wave-number distribution (b) Wave direction distribution

Fig. 7. Wave-number and wave direction distributions for the target, theoretical and measured wave fields of GR for the second experiment.

7. CONCLUSIONS

In this paper we proposed a technique for the measurement of wave fields rendered by loudspeaker arrays with a circular harmonic analysis of the signal captured by a virtual microphone array. The wave field is then analyzed with two evaluation metrics: the first one is based on a RMSE analysis, while the second one performs a wave-number analysis on the captured sound field. The experimental results shows that the proposed technique can be proficiently used for different rendering techniques. If required, the experimental setup can be modified to include also an artificial head for assessing the influence of a listener on the rendered sound field.

8. REFERENCES

[1] A. Berkhout, "Acoustic control by wave field synthesis," *J.Acoust. Soc. Am.*, vol. 93, pp. 2764–2778, 1993.

[2] R. Nicol, "Sound spatialization by higher order ambisonics," in *Proc. 2nd Int. Symp. on Ambisonics and Spherical Acoustics*, 2010.

[3] A. Kuntz and R. Rabenstein, "Wave field analysis using multiple radii measurements," in *proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA '09*, 2009.

[4] T. Abhayapala and A. Gupta, "Non-spherical microphone array structures for 3D beamforming and spherical harmonic analysis," in *proc. of the 11th International Workshop on Acoustic Echo and Noise Control*, 2008.

[5] F. Antonacci, A. Calatroni, A. Canclini, A. Galbiati, A. Sarti, and S. Tubaro, "Soundfield rendering with loudspeaker arrays through multiple beam shaping," in *proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA '09*, 2009.

[6] F. Antonacci, M. Foco, A. Sarti, and S. Tubaro, "Fast tracing of acoustic beams and paths through visibility lookup," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, pp. 812–824, 2008.